AD_____

Award Number DAMD17-98-1-8270

TITLE: Structural Basis of CDK4 Inhibition by p18INK4

PRINCIPAL INVESTIGATOR: Ravichandran Venkataramani

CONTRACTING ORGANIZATION: The Wistar Institute
Philadelphia, Pennsylvania 19104

REPORT DATE: May 1999

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	May 1999	Annual Summary (1 May 98 - 30 Apr 99)

4. TITLE AND SUBTITLE
Structural Basis of CDK4 Inhibition by p18INK4

5. FUNDING NUMBERS
DAMD17-98-1-8270

6. AUTHOR(S)
Ravichandran Venkataramani

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The Wistar Institute
Philadelphia, Pennsylvania 19104

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
This report contains colored photos

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The proteins involved in the G1 to S phase transition have been shown to be particularly important in carcinogenesis. Mainly, the disruptions of the pathways involving the pRb and the p53 tumor suppressors have been shown to be common in cancers. The inactivation of the pRb pathway is commonly achieved either by the mutational inactivation of pRb or INK4 tumor suppressors or by the uncontrolled over-expression of Cyclin D1. Cyclin D1 activates Cyclin Dependent Kinase 4/6 (CDK4/6) which phosphorylate pRb. The INK4 family of proteins are specific inhibitors of Cyclin D1-CDK4/6 complexes. A recently identified protein, $p19^{ARF}$, which is an alternatively spliced transcript of $p16^{INK4a}$, has been shown to stabilize p53 by preventing its degradation in a DNA damage independent manner. This proposal is focussed on the study of INK4 family of proteins, $p19^{ARF}$, and Cyclin D1. We present here the crystal structure of $p18^{INK4c}$ and its implications for CDK4/6 binding. Based on the structure a CDK binding surface is predicted which other researchers have independently confirmed. We also present here the initial characterization of the recombinant cyclin D1 and human $p19^{ARF}$ in *E.coli.*

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer Tumor suppressor, Oncogene, p18INK4c, p19ARF, Cyclin D1	25
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

\ Where copyrighted material is quoted, permission has been obtained to use such material.

\ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

\ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

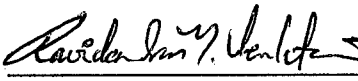
____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

\ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

\ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

\ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.



PI - Signature

May 27th, 1999

Date

TABLE OF CONTENTS

INTRODUCTION

Recent research in the eukaryotic cell cycle has revealed that the mutation of key cell cycle regulatory proteins can lead to tumorigenesis. The proteins involved in the G1 to S phase transition have been shown to be particularly important in carcinogenesis. Mainly, the disruptions of the pathways involving the pRb and the p53 tumor suppressors have been shown to be common in cancers. The inactivation of the pRb pathway is commonly achieved either by the mutational inactivation of pRb or INK4 tumor suppressors or by the uncontrolled over-expression of Cyclin D1. High level of the cyclin D1 oncogene is thought promote uncontrolled cell division by binding and aberrantly elevating the activity of Cyclin Dependent Kinase 4/6 (CDK4/6) which phosphorylate pRb. The INK4 family of proteins are specific inhibitors of Cyclin D1-CDK4/6 complexes [1]. The p53 pathway, on the other hand, is regulated through DNA damage dependent or DNA damage independent mechanisms. A recently identified protein, $p19^{ARF}$, which is an alternatively spliced transcript of $p16^{INK4a}$, has been shown to stabilize p53 by preventing its degradation in a DNA damage independent manner [2]. Moreover, $p19^{ARF}$ is activated by pRb, thus linking the two major tumor surveillance pathways [3]. This proposal is focussed on the study of INK4 family of proteins, $p19^{ARF}$, and Cyclin D1. These studies will provide valuable insights into the structural mechanisms underlying the regulation of the pRb and p53 pathway. The insight gained will be invaluable in the design of small molecule drugs that could be used in the treatment of breast cancer.

RESEARCH ACCOMPLISHMENTS

The specific aims outlined in the approved statement of work are

1. Determine the structure of $p18^{INK4c}$
2. Prepare mutants of $p18^{INK4c}$ defective in CDK4/6 inhibition
3. Determine the X-ray crystal structure of $p18^{INK4c}$ mutants defective in CDK4/6 inhibition
4. Prepare CDK4-$p18^{INK4c}$ and/or CDK6-$p18^{INK4c}$ for complex for structure determination

Of these specific aims, specific aim #1 was successfully completed and the work was published [4]. Another research group [5, 6] accomplished specific aim #4. In the light of their work, specific aims #2 and #3 can be inferred from the crystal structures of the INK4 protein-CDK6 complexes. Thus, we have redirected the focus of our remaining studies to include the study of cyclin D1 and $p19^{ARF}$ proteins. We have described below the progress made so far with the crystal structure determination of $p18^{INK4c}$ and the characterization of cyclin D1 and $p19^{ARF}$.

DETERMINATION OF THE CRYSTAL STRUCTURE OF $P18^{INK4C}$

We have determined the crystal structure of $p18^{INK4c}$ to 1.95 Å resolution [4] and the atomic coordinates have been deposited in the PDB protein structure database (Accession Number: 1IHB). The crystal structure reveals an elongated molecule comprised of five ankyrin repeat units. Each ankyrin repeat contains a beta-strand helix-

turn-helix extended strand beta-strand motif that associates with neighboring motifs through beta-sheet, and helical bundle interactions. A large percentage of residues that are conserved among INK4 proteins play important roles in protein stability [4].

Moreover, analysis of the tumor-derived mutations of $p16^{INK4a}$ when mapped on to the structure of $p18^{INK4c}$ seem to show that they play important roles in protein stability [4]. Some residues were on the surface and yet were associated with $p16^{INK4a}$ tumor derived mutations or strictly conserved amongst the INK4 family of proteins. These residues seem to be clustered on the surface of $p18^{INK4c}$ and were predicted to form the CDK6 binding surface [4].

This prediction has been proven to be correct by the work of Russo et.al and Botherton et.al describing the crystal structures of the complexes of CDK6 with INK4 family members [5, 6]. The comparision of the predicted CDK4 binding surface by the mapping of the strictly conserved residues among the INK4 family members on the surface of $p18^{INK4c}$ (Fig 1A) [4] to the actual residues of $p19^{INK4d}$ involved in CDK6 binding (Fig 1B) [6].

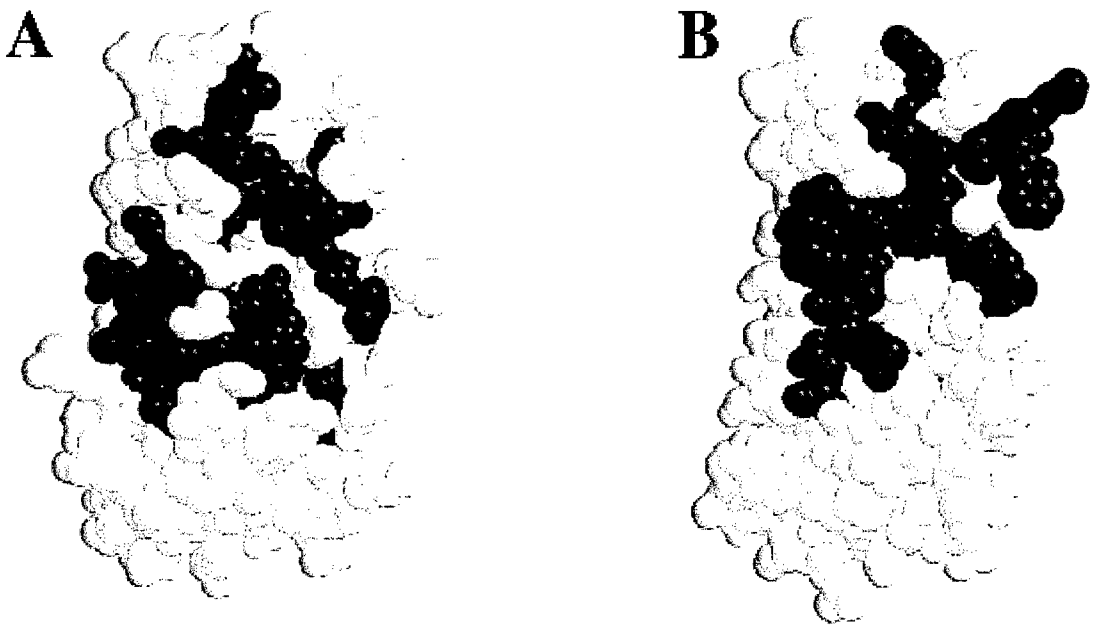


Figure 1. The comparision of the predicted CDK binding surface of $p18^{INK4c}$ **(A)** with the actual CDK6 binding surface of $p19^{INK4d}$ **(B)** reveals the accuracy of the prediction.

CHARACTERIZATION OF RECOMBINANT CYCLIN D1

Previously in our lab we had determined that the recombinant human cyclin D1 is insoluble when overexpressed in BL21 (DE3) *E.Coli* cells (Raymond Trievel, personal communication). Most of our work over the last year has been concentrated at trying to solublize and refold recombinant Cyclin D1 *in vitro* for further studies. We have been successful in solubilizing cyclin D1 in 6M Guanidine HCl buffer. Following solubilization and stepwise dialysis to remove the denaturant has yielded cyclin D1 that is predominantly in the soluble fraction. However, size exclusion chromatography of these cyclin D1 samples reveal that the protein is aggregated and thus not suited for our studies. Moreover, the experiment was repeated under various conditions like high and low salt,

high and low reducing agent concentration and in the presence of organic solvents such as acetonitrile without success.

We modeled the structure of cyclin D1 based on the available structure of cyclin A to serve as a guideline for rational mutations. We inferred from this model that 10 out of the 13 cysteines in cyclin D1 were present on the surface of the protein. We further hypothesized that these cysteines may hinder refolding by forming intermolecular disulfide bonds. Thus, we decided to mutate these cysteines to other amino acids based on homology with cyclin D2 and D3. The wild-type cyclin D1 was then mutated to create a ten-cysteine mutant. This ten-cysteine mutant was found also to be insoluble and to behave very similar to the wild type cyclin D1 when expressed in bacteria. We are presently in the process of expressing wild type and the ten-cysteine mutant of Cyclin D1 using the baculovirus expression system.

CHARACTERIZATION OF HUMAN P19^{ARF}

We have cloned human p19^{ARF} and mouse p19^{ARF} into expression vectors for expression in *E.coli*. We discovered the proteins do not express in *E.coli* even when fused with another protein like Glutathione-S-transferase (GST). We discovered that both the human and mouse p19^{ARF} have unusually high number of rare arginine codons. We hypothesized that the lack of tRNA for these rare arginine codons in *E.coli* prevents the expression of these proteins. We tested this hypothesis on human p19^{ARF} by constructing a completely artificial gene where all codons are optimized for *E.coli* codon usage. This "artificial" human p19^{ARF} (AHARF) does not express alone and we have only been able to achieve expression of the AHARF as a GST-fusion (Fig. 2A shows the induction of GST-AHARF). The fusion protein binds the Glutathione matrix and Ion exchange matrix poorly (Fig 2B). We are in the process of optimizing the binding of the GST-AHARF to these matrices.

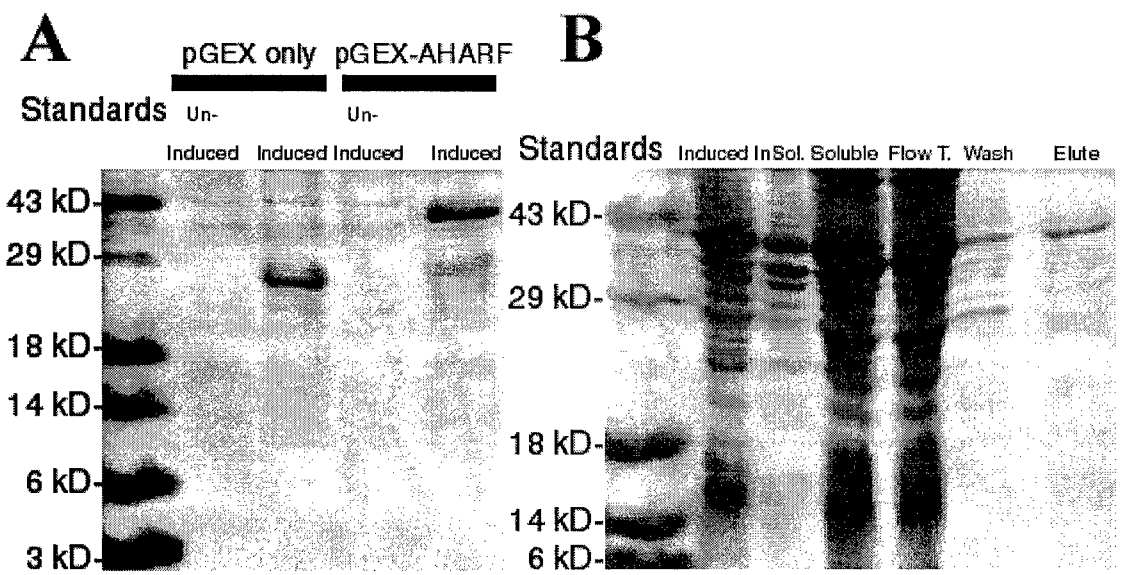


Figure 2: **(A)** The induction of GST-AHARF (pGEX-AHARF, right) when compared with that of GST alone (pGEX, left) clearly shows a insert of about 19kD. **(B)** Samples comparing the Induced, Insoluble (InSol.) and Soluble after sonication of cell pellet, Flow through from the glutathione matrix (Flow T.), Wash and Elute from the

glutathione column. Comparison of the Soluble and Flow through lanes shows that GST-AHARF does not bind the glutathione column completely.

REFERENCES

1. Sherr, C.J. and J.M. Roberts, *Inhibitors of mammalian G1 cyclin dependent kinases.* Genes and Development, 1995. **9**: p. 1149-1163.
2. Sherr, C.J., *Tumor surveillance via the ARF-p53 pathway.* Genes and Development, 1998. **12**: p. 2984-2991.
3. Bates, S., *et al., p14ARF links the tumour suppressors RB and p53.* Nature, 1998. **395**: p. 124-125.
4. Venkataramani, R., K. Swaminathan, and R. Marmorstein, *Crystal structure of the CDK4/6 inhibitory protein p18(INK4c) provides insights into ankyrin-like repeat structure/function and tumor-derived p16(INK4) mutations.* Nature Structural Biology, 1998. **5**(1): p. 74-81.
5. Brotherton, D.H., *et al., Crystal structure of the complex of the cyclin D-dependent kinase Cdk6 bound to the cell-cycle inhibitor p19INK4d.* Nature, 1998. **395**: p. 244-250.
6. Russo, A.A., *et al., Structural basis for inhibition of the cyclin-dependent kinase Cdk6 by the tumour suppressor p16INK4a.* Nature, 1998. **395**: p. 237-243.

LIST OF RESEARCH ACCOMPLISHMENTS

1. **Determination of the structure of $p18^{INK4c}$:** We have determined the structure of $p18^{INK4c}$ to 1.95 Å. The results provided insights into the mechanism of $p18^{INK4c}$ function. The results were published [4] (Attached) and the coordinates were deposited in the PDB Protein Structure Database (Accession Number: 1IHB).
2. **Characterization of recombinant Cyclin D1:** We have characterized cyclin D1 to be insoluble in *E.Coli*. We have further been able to show that cyclin D1 can be refolded using denaturants followed by the stepwise removal of denaturant. But size exclusion chromatography of the refolded cyclin D1 shows that it is aggregated. We are in the process of trying expression of cyclin D1 in baculovirus.
3. **Characterization of recombinant $p19^{ARF}$:** We have expressed recombinant p19ARF in *E.coli*. We are in the process of optimizing the yield of the protein.

LIST OF REPORTABLE OUTCOMES

1. Publication: Venkataramani, R., K. Swaminathan, and R. Marmorstein, *Crystal structure of the CDK4/6 inhibitory protein p18(INK4c) provides insights into ankyrin-like repeat structure/function and tumor-derived p16(INK4) mutations.* Nature Structural Biology, 1998. **5**(1): p. 74-81.
2. Presentation: Marmorstein, R., Venkataramani R., Swaminathan, K., *Structure of the $p18^{INK4c}$ protein: insights into ankryin-like repeat structure/function and CDK4/6 inhibition by the $p16^{INK4}$ tumor suppressor.* 17th International Cancer Congress, Rio de Janeiro, Brazil, 24-28 August 1998.

Crystal structure of the CDK4/6 inhibitory protein p18^{INK4c} provides insights into ankyrin-like repeat structure/function and tumor-derived p16^{INK4} mutations

Ravichandran Venkataramani [1,2], Kunchithapadam Swaminathan[1,4] and Ronen Marmorstein[1,2,3]

p18^{INK4c} is a member of a family of INK4 proteins that function to arrest the G1 to S cell cycle transition by inhibiting the activity of the cyclin-dependent kinases 4 and 6. The X-ray crystal structure of the human p18^{INK4c} protein to a resolution of 1.95 Å reveals an elongated molecule comprised of five contiguous 32- or 33-residue ankyrin-like repeat units. Each ankyrin-like repeat contains a β-strand helix-turn-helix extended strand β-strand motif that associates with neighboring motifs through β-sheet, and helical bundle interactions. Conserved ankyrin-like repeat residues function to facilitate the ankyrin repeat fold and the tertiary interactions between neighboring repeat units. A large percentage of residues that are conserved among INK4 proteins and that map to positions of tumor-derived p16^{INK4} mutations play important roles in protein stability. A subset of these residues suggest an INK4 binding surface for the cyclin-dependent kinases 4 and 6. This surface is centered around a region that shows structural features uncharacteristic of ankyrin-like repeat units.

Progression through the eukaryotic cell cycle correlates with the activity of a family of protein kinases called cyclin-dependent kinases (CDKs), which become fully activated when they are bound to regulatory cyclin subunits and when they are phosphorylated at a particular threonine residue[1]. CDK inhibitory proteins (CKIs) function to inactivate particular CDK–cyclin complexes[2]. Specific CDK–cyclin complexes promote distinct phases of the cell cycle by phosphorylating certain cellular protein substrates, and particular CKIs inhibit distinct phases of the cell cycle.

A major control point for the cell cycle occurs during the G1 to S transition. During this phase of the cell cycle, complexes of CDK4 and CDK6 with D-type cyclins phosphorylate the retinoblastoma gene product, pRb[3]. Upon phosphorylation, pRb stimulates progression into the S phase of the cell cycle by relieving its repressive effect on E2F transcription factors, which function to activate several genes associated with S phase progression[4,5]. The INK4 family of proteins specifically inhibit CDK4–cyclin D1 and CDK6–cyclin D1 complexes and thus play a key role in inhibiting the G1 to S cell cycle transition by preventing pRb-mediated S phase progression[6].

The INK4 protein family includes p16^{INK4} (refs 7,8), p15^{INK4b} (ref. 9), p18^{INK4c} (ref. 10) and p19^{INK4d} (ref. 11). The MTS1 gene encoding p16^{INK4} has received considerable attention since it is deleted, silenced or mutated in a variety of tumor types, including pancreatic adenocarcinoma, esophageal carcinomas, biliary tract cancers, familial melanoma and pancreatic cancer[12–14]. More recently, alterations in the genes encoding the p15^{INK4b} and p19^{INK4d} proteins have also been found in human cancers, but at a much lower frequency than mutations in the MTS1 gene[15,16]. The lack of a correlation between the extent of tumor-derived mutations found in different INK4 proteins with the indistinguishable CDK4/6 inhibitory activity of these proteins, suggests that INK4 proteins may have other non-redundant and/or tissue or cell specific activities.

The homology among the INK4 proteins is restricted to a region containing four 32- or 33-amino acid ankyrin-like repeats initially found in the protein ankyrin[17,18]. To date, more than 150 proteins have been identified that contain ankyrin-like repeat units[19]. Ankyrin repeats are found in many organisms from yeast to humans, and occur in functionally diverse proteins such as enzymes, toxins, and transcription factors[18,19]. Most proteins containing ankyrin-like repeats have at least four copies, and the function of ankyrin repeat-containing proteins is compatible with a role in protein–protein interactions[18,19].

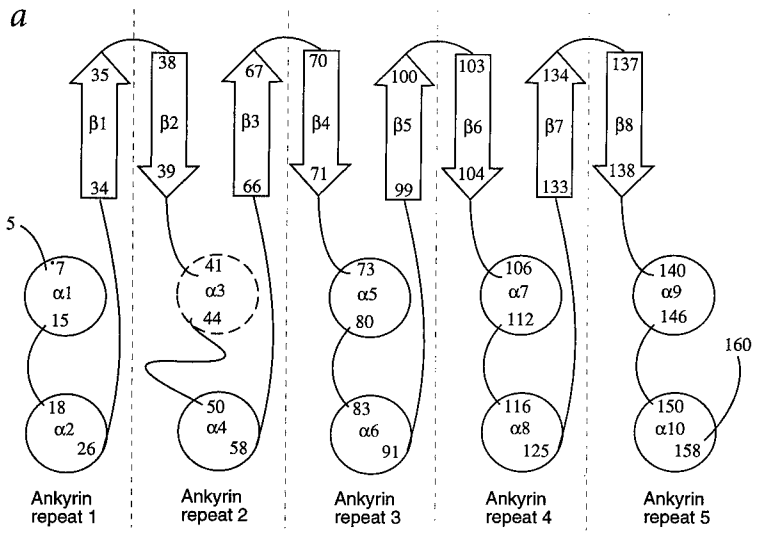
Here we present the 1.95 Å resolution crystal structure of the human INK4 protein, p18^{INK4c}. The structure provides a high resolution framework for understanding ankyrin-like repeat structure and function and affords insights into the structural basis underlying the function of INK4 proteins and tumor-derived mutations in the p16^{INK4} tumor suppressor protein.

Overall structure of the p18^{INK4c} protein

The crystals used for the structure determination contain two p18^{INK4c} protomers in the asymmetric unit cell. The final model contains amino acid residues 5–160 of protomer 1 and residues 7–162 of protomer 2. Both protomers adopt very similar structures, with a root-mean-square (r.m.s) [AUTHOR: OK?] deviation between all atoms of 0.73 Å. Most of the dissimilarity between the two protomers is restricted to the N- and C-terminal regions which are involved in crystal contacts. The p18^{INK4c} protein adopts an elongated shape with dimensions 35 Å x 35 Å x 50 Å, and is comprised of five contiguous 32- or 33-amino acid ankyrin-like segments with similar sequence and structure (Fig. 1). Each repeat

[1]The Wistar Institute, [2]Department of Biochemistry and Biophysics and [3]Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. [4]Current address: Institute of Molecular Agrobiology, National University of Singapore, Singapore, 118240.

Correspondence should be addressed to R.M. *e-mail: marmor@wista.wistar.upenn.edu.*



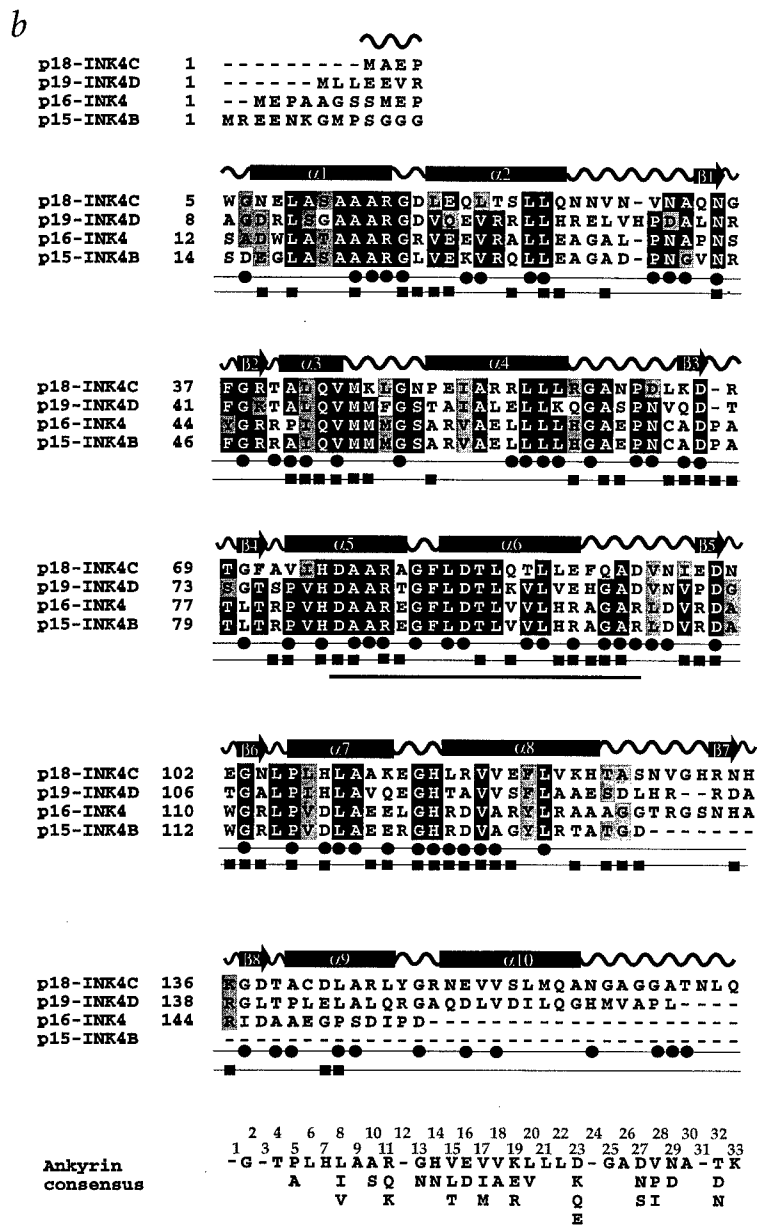


Fig. 1 Secondary structure of the p18^{INK4c} protein and its relationship to the INK4 protein family. ***a***, Topology diagram of the secondary structure elements of the p18^{INK4c} protein depicting the five ankyrin-like repeat units that constitute the structure: α-helices (with helical axis perpendicular to the page) are depicted as circles, and β-strands are indicated by arrows. Residue numbers are indicated at the beginning of each structural element. The uncharacteristically short α3 helix and long loop between α3 and α4 in the second ankyrin repeat is indicated with a broken circle and extended loop respectively. ***b***, Sequence alignment of the INK4 protein family. Black and gray backgrounds indicate identical and conserved residues respectively, found in at least 75% of the proteins at a given position. Secondary structural elements for the p18^{INK4c} protein are indicated above the aligned sequences. Residues that are conserved in ankyrin repeat proteins in two or more homologs are indicated by a black filled circle, and p16^{INK4c} residues that are altered in primary tumors[14,45] are indicated by a black square below the corresponding sequence position. The region corresponding to a 20-residue p16^{INK4}-derived peptide shown to partially mimic the p16INK4-mediated binding and inhibitory activity of CDK4 and CDK6 (ref. 27) is indicated with a solid bar below the sequence. Sequences were aligned with the CLUSTAL program and displayed with the BOXSHADE program. The ankyrin consensus sequence is indicated below with a numbering scheme shown above the corresponding amino acid[18].

unit is comprised of a β-strand helix-turn-helix extended strand β-strand element. The helices and β-strand are arranged in anti-parallel fashion, with the plane of the β-strand regions aligned roughly orthogonal to the axis of the helical segments (Fig. 2*a*). Each ankyrin-like repeat unit interacts with its N- and C-terminal repeat unit through helix–helix and β-sheet interactions, where the helices form helical bundles and the sheet regions form six-residue antiparallel β-hairpins (Fig. 2*b,c*). The only exceptions to these subdomain interactions are found in the first and last ankyrin-like repeat regions, which are each missing terminal β-sheets, presumably because they lack a complementary β-sheet element. In addition, the second ankyrin-like repeat unit has a relatively short first helix (α3) and a long turn region connecting α3 and α4.

Several conserved features characterize and facilitate the fold of the ankyrin-like repeat segments of the 18^{INK4c} structure (Figs 1*b*, 2*b*). A Gly residue at position 2 of a repeat unit always initiates the first β-strand and facilitates a tight turn between neighboring repeats. A Gly residue at position 13 is always found in the turn separating the two helical elements, also facilitating a sharp turn between the two antiparallel helices. A Pro or Ala is usually found at position 5 as the first helical residue, mediating a sharp bend of ~90° between the β-stand and helical segments. An Asp or Asn residue at position 32 also terminates the second β-strand of the repeat unit by hydrogen bonding to the backbone NH of the second residue of the turn that connects to the first β-strand of the proceeding repeat unit.

Extensive van der Waals interactions mediate the association between neighboring ankyrin repeats (Fig. 1*c*). Nearly one-third of the p18^{INK4c} residues are involved in these interactions, which are almost

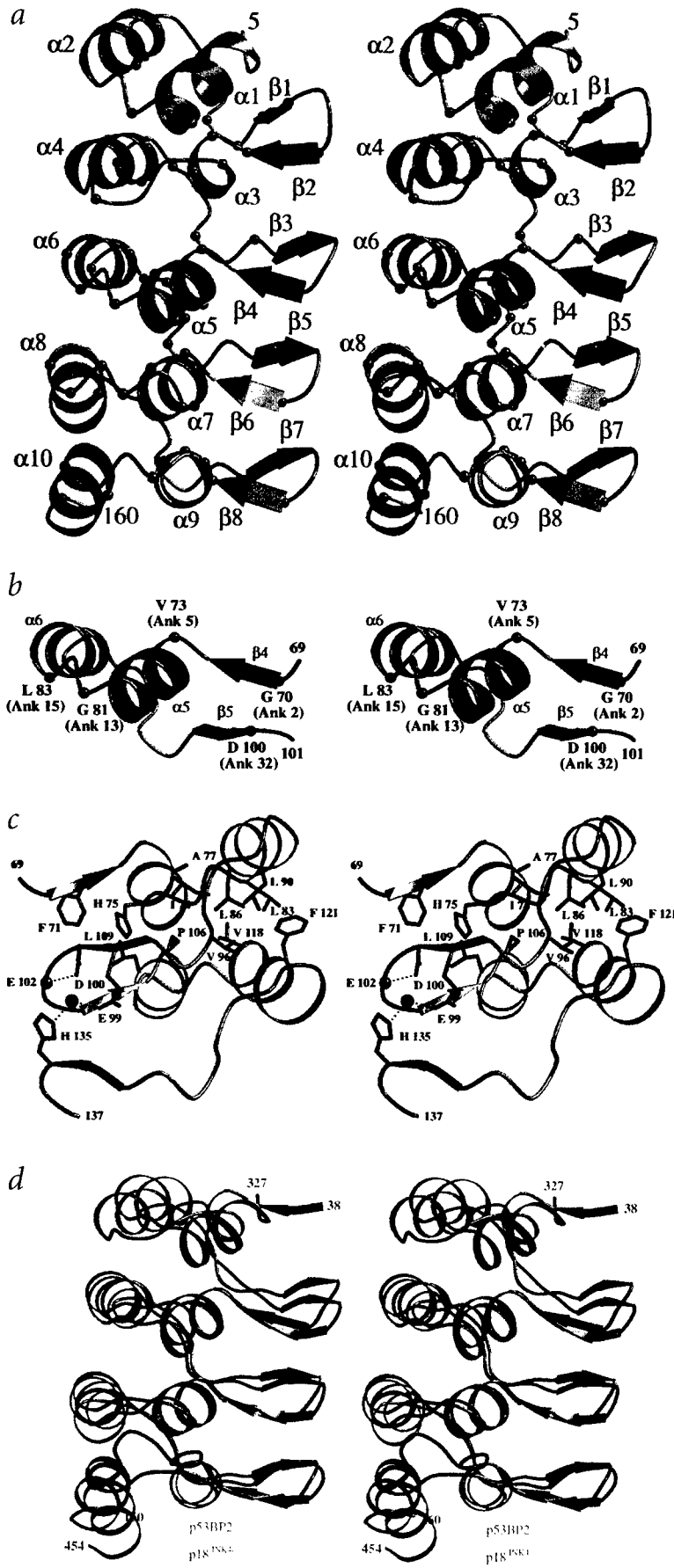


Fig. 2 Structure of the p18^{INK4c} protein and its relationship to the ankyrin-like repeat fold. ***a***, Overall structure of the p18^{INK4c} protein viewed roughly parallel to the helical axis and perpendicular to the plane of the β-sheets. Residues conserved among ankyrin-like domains (red balls and as indicated in Fig. 1*a*) are mapped onto the p18^{INK4c} structure, showing that a large percentage of conserved residues function to stabilize inter-ankyrin repeat interactions. ***b***, Structure of a representative ankyrin-repeat segment of the p18^{INK4c} protein showing the third ankyrin repeat segment and highlighting (red balls) the highly conserved residues between the repeats that facilitate the ankyrin-like repeat fold. The corresponding positions of these residues in the ankyrin consensus sequence are indicated in parentheses. ***c***, Representative inter-ankyrin repeat interactions between ankyrin repeat unit 3 (green) and 4 (aqua). Residues that mediate inter-repeat interactions are highlighted in red. A bound water molecules is indicated with a black ball. The view is rotated by 180° to that shown in (*a*). ***d***, Cα atom superposition of ankyrin-like repeats 1–4 of the p18^{INK4c} protein with the ankyrin repeat region of the 53BP2 protein. Figs 2 and 3 were prepared with the programs MOLSCRIPT[46] and RASTER3D[47].

all localized to the helical segments, thereby stabilizing the helical bundle interactions of the protein. Positions 5, 10, 17, 20, and 21 (numbering is according to ankyrin consensus shown in Fig. 1*b*) of each repeat have conserved hydrophobic residues (usually small) that carry out conserved inter-repeat stabilizing functions. Additional hydrophobic residues also stabilize inter-repeat interactions, but these residues are generally not conserved among the repeat units. The solvent-exposed regions of the helical segments from the first and last repeat regions are generally less hydrophobic than their interior helical counterparts.

The β-hairpins at the junctions of each of the repeat units interact with neighboring β-hairpins predominantly through salt bridges. Except for a recurring water-mediated hydrogen bond between backbone atoms of residues at the tip of each β-hairpin with neighboring hairpins, these salt bridges are mediated by residues that are not generally conserved between repeats. Taken together, the overall fold and elongated shape of the p18^{INK4c} protein appears to depend on the interactions between neighboring ankyrin-repeat units.

Implications for ankyrin-like repeat structure

Correlation of the ankyrin-like repeat consensus sequence with the p18^{INK4c} structure reveals that the residues facilitating the fold of the ankyrin-like repeat segments and mediating interactions between neighboring repeat units are highly conserved in ankyrin repeat-containing proteins (Fig. 1*b*). For example, Gly residues are strictly conserved at positions 2 and 25 of the ankyrin repeat consensus sequence; in the p18^{INK4c} structure, these residues play important roles in facilitating sharp turns between antiparallel β-sheet and helical segments respectively. A Gly residue is also highly preferred at position 13, where it functions to mediate a sharp bend between antiparallel helices of each repeat unit in the p18^{INK4c} structure. In addition, the ankyrin consensus sequence shows that small

hydrophobic residues are preferred at positions 5, 10, 17, 20 and 21, and the p18^{INK4c} structure shows that each of these residues plays an important role in mediating repeat–repeat interactions through helical bundle contacts.

Other residues that appear to be highly preferred in ankyrin-like sequences are Thr at position 4, Leu at positions 6 and 22, His at position 7, and Ala at positions 9, 26 and 30 (Fig. 1*b*). Although these residues are conserved in selected repeat units of the p18^{INK4c} protein, they are not uniformly conserved in each of the repeat units. This suggests that ankyrin repeat structures tolerate small deviations from the consensus sequence and explains why the ankyrin consensus sequence is generally poor as compared to other consensus sequences, such the bacterial helix-turn-helix transcription factors[20] and from the mammalian cyclin-dependent kinases[21]. Together, our observations suggest that the ankyrin repeat unit maintains a conserved three-dimensional fold within other ankyrin repeat-containing proteins. Moreover, the high conservation of hydrophobic residues that appear to facilitate repeat–repeat interactions correlates well with the observation that ankyrin repeat proteins usually contain multiple ankyrin-repeat units (more than four), and suggests that multiple units will align in a very homologous way in other ankyrin repeat-bearing proteins.

The overall structure of the p18^{INK4c} protein appears similar to the recently published p19^{INK4d} solution structure[22], although a more detailed comparison must await the availability of atomic coordinates for the p19^{INK4d} protein. The only other three-dimensional coordinates available for an ankyrin repeat structure is the X-ray crystal structure of the ankyrin repeat-containing domain of 53BP2 bound to the DNA binding region of the p53 protein[23]. The ankyrin-like repeat units and the overall multi-repeat domain from the 53BP2 protein superimposes well with the respective regions of the p18^{INK4c} structure. A superposition of any of the repeat units of the 53BP2 structure with any of the repeat units of the p18^{INK4c} structure (except repeat unit 2 of p18^{INK4c} and repeat unit 4 of 53BP2) shows an r.m.s. deviation between Cα atoms of less than 0.80 Å. Superposition of the entire ankyrin-like domain of the 53BP2 structure with the last 4 repeat units of the p18^{INK4c} structure shows an r.m.s. deviation of Cα atoms of 6.0 Å (Fig. 2*d*). However, most of the differences between the two structures are localized to the regions that deviate from the ankyrin-like repeat structure. These deviations are found in the α3 helix and the proceeding loop region of p18^{INK4c} and in the first helix of the fourth repeat unit of 53BP2. Superposition of Cα atoms with the exclusion of these regions results in an r.m.s. deviation of less than 1.6 Å between the two structures.

In addition to the above noted difference between the ankyrin-like repeat units of 53BP2 and p18^{INK4c} there are differences in the relative arrangement of the β-hairpin segments connecting the ankyrin like repeats. Most notably, in the 53BP2 structure the β-hairpin segments are spaced in a way that allows a continuous antiparallel β-sheet network to be formed throughout the length of the protein, while in the p18^{INK4c} structure (except for one backbone hydrogen bond between the second and third β-hairpin segments) only inter-repeat β-sheet interactions are observed. The p18^{INK4c} structure also shows a recurrent water-mediated interaction between β-hairpin segments that link repeat units; this is not observed in the ankyrin-like repeat units of 53BP2[23].

Structural characterization of two other INK4 proteins, p16^{INK4} (ref. 24) and p19^{INK4d} (ref. 25), has been carried out using solution NMR techniques. These studies confirm that the INK4 proteins are highly helical, with each of the four homologous ankyrin repeat regions containing a helix-turn-helix motif. Aside from the absence of significant β-sheet secondary structure in the NMR characterizations of p16^{INK4} and p19^{INK4d}, the secondary structural elements of the p16^{INK4} and p19^{INK4d} proteins are consistent with the secondary structure of the 53BP2 (ref. 23) and p18^{INK4c} proteins determined by crystallography. Moreover, the recently published solution structure of the p19^{INK4d} protein[22] shows similar secondary and tertiary structure to the 53BP2 and p18^{INK4c} proteins. Also more recently, an NMR secondary structure assignment of the ankyrin-like repeat-bearing protein myotropin was completed; this structure also reveals a helix-turn-helix motif in two contiguous ankyrin-like repeat segments[26].

Together, the structural information obtained for the p18^{INK4c}, p16^{INK4}, p19^{INK4d}, 53BP2, and myotropin ankyrin repeat-containing proteins suggests that ankyrin repeat regions from otherwise dissimilar proteins adopt similar three-dimensional folds. The p18^{INK4c}, 53BP2 and p19^{INK4d} structures, and the relatively high degree of sequence homology within the buried helical regions of ankyrin-like domains (Fig. 1), suggests that the overall three-dimensional fold of multi-repeat ankyrin-domain proteins are in large part maintained by the helical bundle interactions between neighboring repeat units.

Despite the similarity in the three-dimensional scaffold of multi-repeat ankyrin-domain proteins, the relative lack of sequence homology within the solvent exposed helical regions, the β-hairpin regions, and the loop regions connecting helical elements to the second β-sheet within each ankyrin-like repeat, implicates these regions as likely protein–protein interaction sites (Fig. 2*a*). Indeed, the relatively non-globular shape and large exposed surface of the ankyrin-repeat 'super-structure' is consistent with a role in mediating protein–protein interactions, a common feature among proteins containing ankyrin-like repeats[19]. The β-hairpin region in particular shows the greatest sequence divergence and structural variability (between 53BP2 and p18^{INK4c}) than other regions of the ankyrin repeats, suggesting that the protein-specific binding activity of particular ankyrin repeat-containing proteins may employ the β-hairpin regions of the repeats. This prediction is supported by the observation that the ankyrin domain of 53BP2 interacts with the p53 protein through a β-hairpin region within the fourth ankyrin-repeat region of 53BP2 (ref. 23).

Implications for INK4 proteins and p16^{INK4} mutations

Sequence alignment of the INK4 proteins p16^{INK4} (ref. 7,8), p15^{INK4b} (ref. 9), p18^{INK4c} (ref. 10), and p19^{INK4d} (ref. 11) reveals a high degree of homology, with 34% identity and 53% similarity overall within the first four ankyrin-like repeat regions of the proteins (Fig. 1*b*). p18^{INK4c} and p19^{INK4d} contain a fifth ankyrin-like domain and show 21% sequence identity and 51% similarity within this region (Fig. 1*b*). The p18^{INK4c} structure shows that nearly 80% of the residues conserved among the INK4 proteins are associated with facilitating complementary interactions between neighboring repeat units. Of these, 80% are generally conserved in ankyrin-like repeat sequences (Fig. 1*b*). This observation suggests that the INK4 proteins have very similar three-dimensional structures. Based on that assumption, we mapped the p16^{INK4} tumor-derived mutations onto the p18^{INK4c} structure (Fig. 3*b*). Most of the mutations localize to residues involved in ankyrin repeat conformation or repeat–repeat interactions, suggesting that a large percentage of p16^{INK4} missense mutations result in reduced p16^{INK4} stability with possible unfolding of the

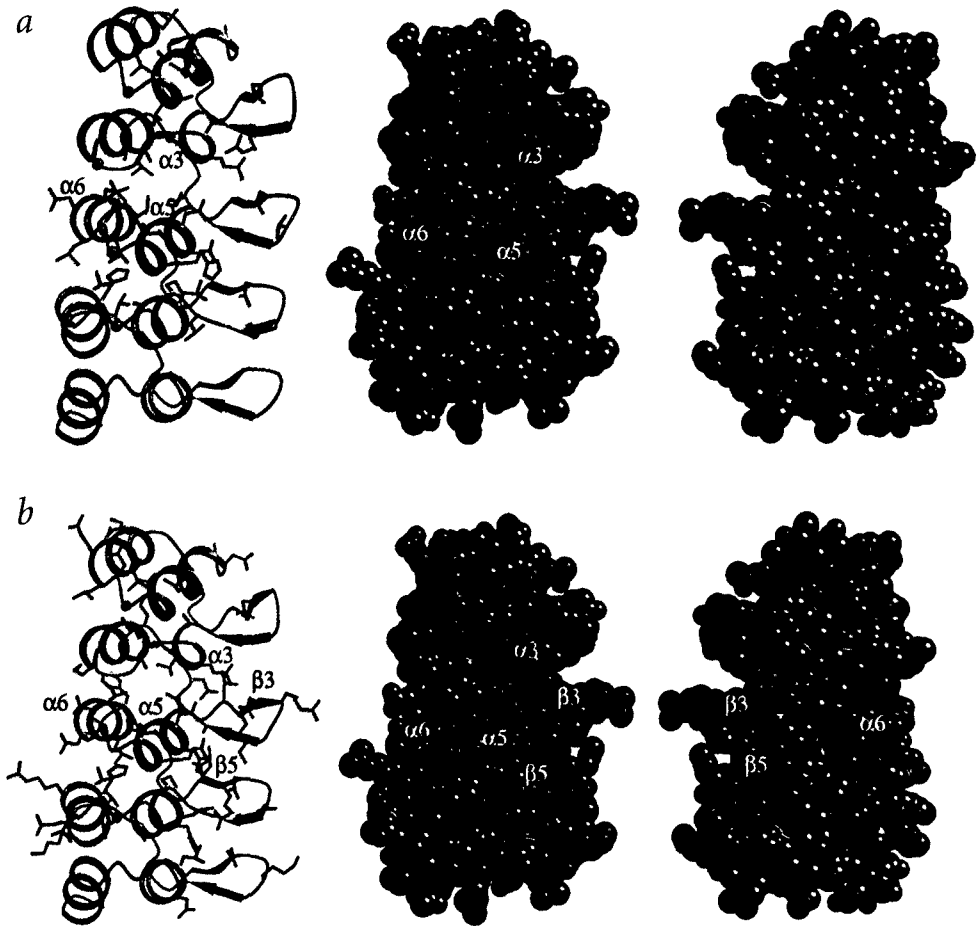


Fig. 3. Implications of the p18^{INK4c} structure for INK4 proteins and tumor-derived p16^{INK4} mutations. ***a***, Residues that are strictly conserved among INK4 proteins are highlighted in red; left, conserved residues are superimposed onto a schematic representation of the p18^{INK4c} structure; center, a space filling model in the same orientation as the left view; right, a space filling model shown 180° from the orientation shown in the center and left views. ***b***, Residues of p18^{INK4c} that correspond to positions of p16^{INK4} tumor-derived mutations are highlighted in red. The views are as indicated in (*a*).

protein. This suggestion is consistent with studies indicating that the p16^{INK4} tumor-derived mutations commonly found in familial melanoma (P114L, G101W, and D84H) show reduced protein stability *in vitro*[24].

Solvent exposed residues in the nascent p18^{INK4c} molecule that are strictly conserved among INK4 proteins are clustered around the α3 helix and the proceeding loop of ankyrin-like repeat unit 2, and the α5-turn-α6 region of ankyrin-like repeat unit 3 (Fig. 3*a*). Interestingly, solvent-exposed tumor-derived p16^{INK4} mutations also cluster to these regions, and to the β3 and β5 strands flanking the α3 and α5 helices respectively (Fig. 3*b*). Together, the high degree of conservation and mutational sensitivity of the α3-loop, and α5-turn-α6 regions of p18^{INK4c} suggest that these regions play an important role in CDK4 and CDK6 inhibition. Specifically, Glu 43, Met 45, Asp 76, Ala 77, Arg 79, Gly 81 and Thr 85 are solvent exposed residues that are strictly conserved and mutationally sensitive p16^{INK4} hotspots, thus implicating them in forming a contact surface for cyclin-dependent kinases 4 and 6.

Fahraeus and coworkers investigated the ability of overlapping 20-residue peptides spanning the p16^{INK4} amino-acid sequence to bind CDK4 and CDK6 (CDK4/6), and to inhibit CDK4/6-mediated phosphorylation of pRb *in vitro*[27]. The results of these studies are in agreement with our findings that the α5-turn-α6 region of p18^{INK4c} is important for inhibition of CDK4/6. Specifically, a 20-residue peptide encoding residues 84–103 of p16^{INK4}, and corresponding to the C-terminal half of the α5-helix through the α6-helix of p18^{INK4c} (Fig. 1*b*), bound CDK4/6 more avidly than other peptides tested. This peptide also inhibited CDK4/6-mediated phophorylation of pRb more potently then two control peptides that showed negligible CDK4/6 binding activity. Moreover, an Asp to Ala substitution at position 92 of this peptide (corresponding to Asp 84 in p18^{INK4c}), predicted from our structure to be located on the protein surface and therefore to be available for protein interaction, increased CDK4/6 binding and CDK4/6-mediated inhibition of pRb. A Leu to Ala substitution at position 94 of this peptide (corresponding to Leu 84 in p18^{INK4c}), predicted form our structure to be buried and to thereby be important for protein stability, decreased both CDK4/6 binding and CDK4/6-mediated inhibition of pRb. Interestingly, a 20-residue peptide harboring a p16^{INK4} sequence corresponding to the α3-loop region of p18^{INK4c} (residues 32–51 in p18^{INK4c}) also had partial CDK4/6 binding activity.

The α3 helix of the second repeat unit has an unusual structural feature relative to the other repeats in the p18^{INK4c} protein; that is, the first helix of the helix-turn-helix submotif is relatively short (only one turn) and thereby contains an uncharacteristically long turn region between helices α3 and α4 (Figs 2*a*, 4*a*). As a result, the helical bundle interactions between the α3 and α5 helices are less extensive than the other helical bundle interactions in the p18^{INK4c} structure. Thus, the junction between repeat units 2 and 3 is more open than the corresponding junctions of neighboring repeat units, resulting in a small cavity in this region of the protein (Fig. 4*b*). The unusual feature of the α3 helix was also noted in the recently published three-dimensional solution structure of p19^{INK4d} (ref. 22), and in the NMR secondary structure characterization studies of p19^{INK4d} (ref. 25) and p16^{INK4} (ref.

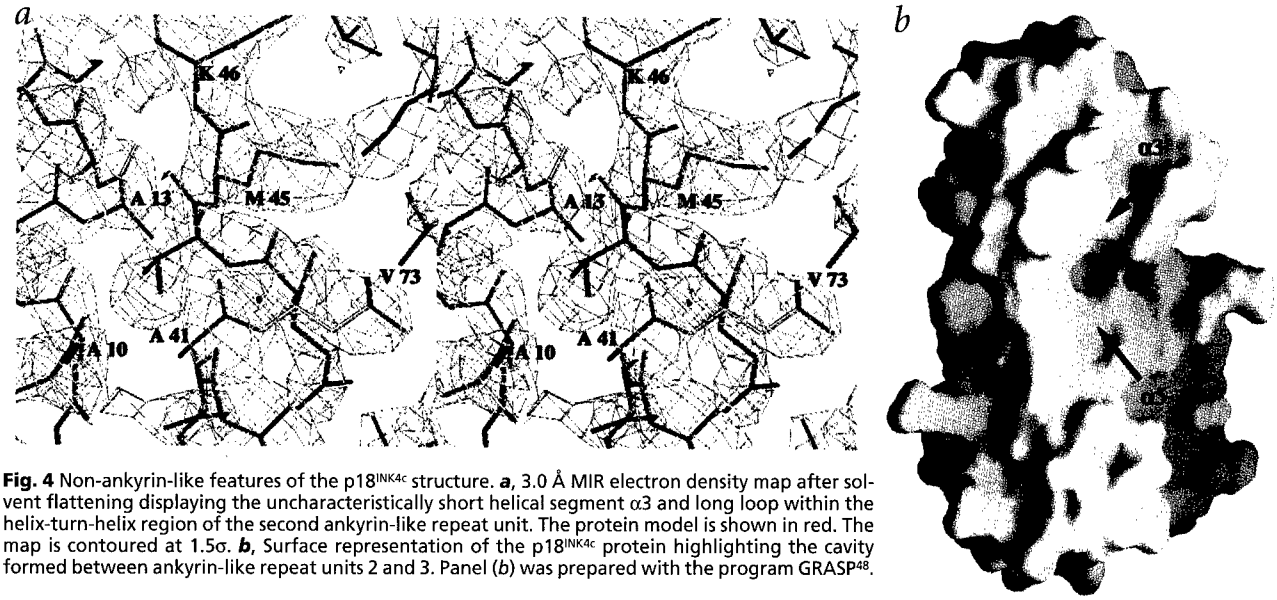


Fig. 4 Non-ankyrin-like features of the p18^{INK4c} structure. ***a***, 3.0 Å MIR electron density map after solvent flattening displaying the uncharacteristically short helical segment α3 and long loop within the helix-turn-helix region of the second ankyrin-like repeat unit. The protein model is shown in red. The map is contoured at 1.5σ. ***b***, Surface representation of the p18^{INK4c} protein highlighting the cavity formed between ankyrin-like repeat units 2 and 3. Panel (*b*) was prepared with the program GRASP[48].

24), although in the secondary structure characterization studies the α3 region is non-helical. Thus, an unusual deviation of the second ankyrin-like repeat unit from the general ankyrin repeat fold does appear to be a characteristic feature of the INK4 proteins. This deviation, coupled with the high degree of sequence conservation and mutational sensitivity of the α3-loop and α5-turn-α6 regions of p18^{INK4c}, further implicates this region to play an important role in inhibition of CDK4 and CDK6.

Coleman and coworkers[28] have used site-directed mutagenesis coupled with a two-hybrid binding assay to identify residues in CDK4 that may be involved in p16^{INK4} binding. Of the single site mutations they tested, two showed an activity at least 20-fold reduced from wild type: K22A, and R24A. In addition, three single-site mutations showed binding activity that was two to three fold greater than wild type levels: D25A, Y17F, and T172A. Wolfel and coworkers[29] have also isolated an R24C mutation in CDK4 that renders it unable to bind p16^{INK4} and therefore resistant to p16^{INK4} inhibition. Assuming that the CDK4 structure is similar to the homologous CDK2 structure that has been determined by X-ray crystallography[30], each of these mutationally sensitive residues (except Thr 172) would map to the N-terminal lobe of the bi-lobed CDK4 protein. The implication that p16^{INK4} interacts with the N-terminal lobe of CDK4 is reminiscent of the way in which the p27^{Kip1} cell-cycle inhibitory protein binds CDK2[31]. However, the fact that CDK inhibition by p16^{INK4} is specific for CDK4 and CDK6 and that the R24C mutation isolated by Wolfel and coworkers is insensitive to inhibition by p27^{Kip1}, suggests that p16^{INK4}, and other INK4 proteins, inhibits CDK4/6 by a distinct mechanism.

Together, the results discussed above suggests that the α3-loop and α5-turn-α6 regions of INK4 proteins interact with the N-terminal lobe region of CDK4 and CDK6. Presumably, these interactions facilitate inhibition either through displacement of the cyclin D regulatory subunit or through direct inhibition of the CDK catalytic site located at the interface of the N- and C-terminal lobes, or through a combination of both mechanisms. Consistent with the first mechanism, Coleman and coworkers have found that all CDK4 mutations that perturb p16^{INK4} binding also perturbed cyclin D1 binding[28]. Although the high-resolution structure of the p18^{INK4c} protein provides insights into INK4-mediated inhibition of CDK4 and CDK6 and the molecular basis underlying p16^{INK4} tumor-derived mutations, a more complete understanding awaits the structure of the complex between an INK4 protein and CDK4 or CDK6.

Methods

Protein expression and purification. The DNA segment encoding the full-length p18^{INK4c} protein was amplified from the plasmid pUC-p18C[10] (provided by Y. Xiong, University of North Carolina) using polymerase chain reaction (PCR) with pairs of oligonucleotide primers in which p18^{INK4c} specific primers were linked to upstream *Nde*I and downstream *Bam*HI sites. This amplified fragment was subcloned into the PRSETA T7-polymerase based expression vector and transformed into *Escherichia coli* strain DH5 to facilitate plasmid isolation and dideoxynucleotide sequencing (Wistar DNA Core facility) of the amplified p18^{INK4c} gene. For protein over-production, the PRSETA/p18^{INK4c} expression plasmid, confirmed to harbor the native p18^{INK4c} sequence, was transformed into *E. coli* strain BL21(DE3).

Recombinant p18^{INK4c} protein was produced by growing p18^{INK4c}-transformed BL21(DE3) cells at 20 °C to an A_{595} of 0.4 and inducing with 1 mM isopropyl-1-thio-D-galactopyranoside (IPTG) to an A_{595} of ~1.0. The collected cells were lysed by sonication in a low-salt buffer (LSB) containing 40 mM Tris (pH 8.5), 50 mM NaCl, 50 mM $(NH_4)_2HPO_4$, 1.0 mM dithiothreitol (DTT), and 0.1 mg ml^{-1} phenylmethylsulfonyl fluoride (PMSF). The supernatant was isolated by centrifugation and applied directly to a Q-sepharose fast flow column (Pharmacia). Fractionation was achieved using a 0.05–0.4 M NaCl gradient in LSB, and peak fractions were pooled and concentrated to ~20 mg ml^{-1} by ultracentrifugation using a Centriprep-10 microconcentrator (Amicon). p18^{INK4c} was further purified by gel filtration FPLC using a Superdex-75 column (Pharmacia) in LSB buffer. Peak fractions were concentrated to ~15 mg ml^{-1} using a Centricon-10 microconcentrator (Amicon). Protein aliquots were flash-frozen at -70 °C and thawed prior to crystallization as needed. Protein purity was judged to be greater than 95% by SDS-gel electrophoresis and Coomassie blue staining.

Selenomethionine-derivatized p18^{INK4c} protein was prepared by growing pRSET/p18^{INK4c}-transformed *E. coli* strain B834 (DE3) (Novagen) in MOPS-based minimal media supplemented with 50 mg l^{-1} L-selenomethionine and other amino acids at the suggested concentrations[32]. Cells were grown at 37°C to an A_{595} of 0.4 and induced with 1 mM IPTG to an A_{595} of ~1.0. The p18^{INK4c} protein was

isolated from inclusion bodies by suspending the cell pellet in LSB buffer supplemented with 6 M urea, 10 mM DTT and 20 mM EDTA. Urea was removed by stepwise dialysis at 4°C. Soluble selenomethionine-derivatized $p18^{INK4c}$ protein was concentrated by ultracentrifugation and further purified by anion-exchange and gel filtration chromatography as described above except for the inclusion of 10 mM DTT and 20 mM EDTA throughout the protein purification to prevent oxidation of the derivatized protein[33]. Quantitative amino acid analysis of selenomethionine-derivatized $p18^{INK4c}$ protein confirmed that more than 90% of the methionine residues had been replaced. Aliquots of the protein were flash-frozen at -70 °C and thawed prior to crystallization as needed.

Crystallization and data collection. Crystals of the $p18^{INK4c}$ protein were obtained by vapor diffusion at 4 °C, using 2 µl hanging drops containing 5 mg ml^{-1} protein, 20 mM Tris (pH 8.5), 25 mM $(NH_4)_2HPO_4$, 0.5 mM DTT, 7% PEG 6000 (polyethylene glycol, average molecular mass 6000 M_r), and 1 M NaCl equilibrated over a reservoir containing 14% PEG 6000 and 2 M NaCl. Crystals grew in several days to a typical size of 0.4 mm × 0.2 mm × 0.2 mm in the space group $P2_12_12$, with a = 55.6, b = 151.6, c = 40.5 and $\alpha = \beta = \gamma = 90°$, and contain two protomers per asymmetric unit with ~45% solvent. Crystals were transferred to a harvest solution (HS) containing 20 mM Tris (pH 8.5), 25 mM $(NH_4)_2HPO_4$, 15% PEG 6000 and 2 M NaCl, transferred stepwise to HS supplemented with 25% glycerol, and frozen in liquid nitrogen-cooled liquid propane prior to data collection.

All data were collected at 110 K on either a Siemens X-100A multiwire area detector with CuKα radiation from a Rigaku RU-200 generator with Franks mirrors, a R-AXIS II image plate detector with CuKα radiation from a Rigaku RU-200 generator with MSC/Yale mirrors, or a MAR image plate detector equipped with a Siemens rotating anode generator with MSC/Yale mirrors. Data obtained from the X-100A detector were processed with XDS[34,35], and data obtained with the R-AXIS II image plate and MAR detectors were processed with DENZO[36].

Structure determination and refinement. The structure was determined to 3.0 Å resolution by multiple isomorphous replacement (MIR) using phase information from three heavy atom derivatives: $K_2Pt(CN)_4$ (5 mM, 24 h soak), $HgCN_2$ (5 mM, 24 h soak) and selenomethionine-derivatized $p18^{INK4c}$. Two heavy atom sites from the $HgCN_2$ derivative were easily identified from difference Patterson maps, and the positions of the remaining two derivatives were identified using difference Fourier techniques. Heavy atom parameters were refined with the program PHASES[37] (Table 1). The MIR map was further improved by applying the solvent flattening and non-crystallographic two-fold averaging procedures implemented in the program PHASES[37]. This map permitted clear interpretation of the secondary structure elements and many of the buried side chains. The programs TOMFRODO[38] and O[39] were used to build a partial Cα backbone for one of the two protomers in the asymmetric unit cell using the alanized structure of the ankryin repeat regions extracted from the 53BP2 structure (provided by N. P. Pavletich, Memorial Sloan-Kettering Cancer Center)[23] as a guide. The positions of the two ordered selenomethionine positions and the cystine-bound mercury position were used to register the placement of amino acid side chains into the electron density map. Once the first protomer was adjusted to density, the second protomer was generated using a symmetry operation defined by the eight non-crystallographically related heavy atom positions. The experimental map allowed for the placement of ~80% of the protein atoms.

Model refinement was initially carried out to 3.0 Å resolution using native data set 1 (Table 1). Before refinement, a randomly selected 10% of the data were omitted from the refinement and used as a 'free data set' to monitor subsequent calculations[40], and the model was refined against the remaining 90% of the data (working data set). Conventional position refinement was initially carried out with non-crystallographic (NCS) restraints imposed between the two protomers in the asymmetric unit with the X-PLOR program[41]. After adjustment of the model as indicated by Sigmaa-weighted $2\boldsymbol{F}_o - \boldsymbol{F}_c$ and $\boldsymbol{F}_o - \boldsymbol{F}_c$ difference maps with the program TOMFRODO[38], O[39] and MAIN (D. Turk, Ph.D. thesis, TU Muenchen), the model was refined at 2.5 Å without NCS restraints and using simulated annealing with X-PLOR[42]. Subsequent refinements employed an improved native data set (native data 2, Table 1). Iterative cycles of simulated annealing refinement with X-PLOR[42] and manual model building using Sigmaa-weighted $2\boldsymbol{F}_o - \boldsymbol{F}_c$ and $\boldsymbol{F}_o - \boldsymbol{F}_c$ difference maps were extended in steps to

Table 1 Structure statistics

			Derivatives		
	Native	Native 2	$K_2Pt(CN)_4$ (5 mM, 24 h)	$HgCN_2$ (5 mM, 24 h)	Selenomethionine derivative
Resolution (Å)	15–2.5	15–1.95	15–3.0	15–2.5	15–3.0
Completeness (%)	94.9	95.0	71.5	83.7	92.1
R_{symm} (%)	0.074	0.042	0.083	0.074	0.046
R_{merge}[1] (%)			0.151	0.106	0.056
Phasing Power[2] (10–3 Å)			1.72	2.17	1.73
Cullis R-factor[3]			0.529	0.479	0.535
Overall figure of merit				0.623 (10–3 Å)	

Refinement statistics[4]

		Atoms			
Resolution	Reflections	Protein	Water	$R_{working}$[3]	R_{free}[5]
8–1.95 Å		2,383	213		
all F	22,215			21.1%	29.1%
$F > 3\sigma$	21,441			20.5%	28.4%
	Bond lengths	Bond angles	B-values[6]	2-molecules[7]	
R.m.s.d.	0.009 Å	1.5°	2.60 Å^2	0.734 Å	

[1] $R_{merge} = \Sigma_{hkl} \Sigma_i |I_{hkl,i} - \langle I_{hkl} \rangle| / \Sigma_{hkl} \Sigma_i I_{hkl,i}$.
[2] Phasing power = $[\Sigma_{hkl} \boldsymbol{F}_h^2 / \Sigma_{hkl} (\boldsymbol{F}_{PH,obs} - \boldsymbol{F}_{PH,calc})^2]^{1/2}$.
[3] R-factor = $\Sigma_{hkl} ||\boldsymbol{F}_{PH}| - \boldsymbol{F}_{H,calc}| / \Sigma_{hkl} |\boldsymbol{F}_{PH}|$, for centric reflections $\Sigma_{hkl} |\boldsymbol{F}_{obs} - \boldsymbol{F}_{calc}| / \Sigma_{hkl} \boldsymbol{F}_{obs}$
[4] Model contains residues 5–160 of protomer 1 and residues 7–162 of protomer 2 (out of 169 residues).
[5] R_{free} is calculated for 10% of the data.
[6] Atomic B-values are for bonded atoms.
[7] Value is for all atoms of residues 7–160 of promoters 1 and 2 in the asymmetric unit cell.

resolution limits of 2.5 Å, 2.25 Å, 2.0 Å and 1.95 Å. After application of a bulk solvent correction[43], tightly constrained atomic *B*-factors were adjusted, and water molecules were built into regions that showed strong F_o - F_c peaks and made stereochemically feasible hydrogen bonds. The correctness of the model was checked against simulated annealing omit maps[44] over the entire structure by omitting 15 residues at a time, and the model was adjusted appropriately. A last round of refinement resulted in a model with good geometry (r.m.s.d. bond length = 0.009 Å, r.m.s.d. bond angle = 1.5°) and a working *R*-factor of 21.1% with a free *R*-factor of 29.1% using all reflections between 8 and 1.95 Å (Table 1). The final model includes residues 5–160 of protomer 1 and residues 7–162 of protomer 2. A Ramachandran plot showed no residues in disallowed regions.

Coordinates. Coordinates of the p18^{INK4c} protein have been deposited in the Brookhaven Protein Data Bank (accession code 1IHB). They can also be obtained from the corresponding author by request.

Acknowledgments

We thank N.P. Pavletich (Cellular Biochemistry and Biophysics Program, Sloan-Kettering Cancer Center) for providing coordinates of 53BP2; Y. Xiong (Department of Biochemistry and Biophysics, the University of North Carolina) for providing the gene encoding p18^{INK4c}; M. Lewis and G. Van Duyne and members of their laboratory (Department of Biochemistry and Biophysics, the University of Pennsylvania Medical School) for use of their MAR image plate detector system; D. Christianson and members of his laboratory (Department of Chemistry, the University of Pennsylvania) for use of their R-AXIS II image plate detector system; and G. Van Duyne, M. Lewis, P. Loll, R. Burnett, D. King, X. Li, Y. Mo, S. Stayrook, T. Stams, C. Lesburg, S. Benson, J. Rux and J. Taylor for useful discussions. This work was supported by a grant from the W. W. Smith Charitable Trust to R. M.

Received 12 November 1997; accepted 5 December 1997

1. Morgan, D.O. Principles of CDK regulation. *Nature* **374**, 131–134 (1995).
2. Elledge, S.J. & Harper, J.W. Cdk inhibitors: on the threshold of checkpoints and development. *Curr. Opin. Cell Biol.* **6**, 847–852 (1994).
3. Hollingsworth, R.E., Phang-Lang, J. & Lee, W.-H. Integration of cell cycle control with transcriptional regulation by the retinoblastoma protein. *Curr. Opin. Cell Biol.* **5**, 194–200 (1993).
4. Weinberg, R.A. The retinoblastoma protein and cell cycle control. *Cell* **81**, 323–330 (1995).
5. Chow, K.N.B., Starostik, P. & Dean, D.C. The Rb family contains a conserved cyclin-dependent-kinase-regulated transcriptional repressor motif. *Mol. Cell. Biol.* **16**, 7173–7181 (1996).
6. Ranade, K., *et al.* Mutations associated with familial melanoma impair p16^{INK4} function. *Nature Genet.* **10**, 114–116 (1995).
7. Serrano, M., Hannon, G.J. & Beach, D. A new regulatory motif in cell cycle control causing specific inhibition of cyclin D/CDK4. *Nature* **366**, 704–707 (1993).
8. Kamb, A., *et al.* A cell cycle regulator potentially involved in the genesis of many tumor types. *Science* **264**, 436–440 (1994).
9. Hannon, G.J. & Beach, D. p15^{INK4B} is a potential effector of TGF-—induced cell cycle arrest. *Nature* **371**, 257–261 (1994).
10. Guan, K.-L., *et al.* Growth suppression by p18, a p16$^{INK4/MTS1}$- and p14$^{INK4B/MTS2}$-related CDK6 inhibitor, correlates with wild-type pRb function. *Genes Dev.* **8**, 2939–2952 (1994).
11. Chan, F.K.M., Zhang, J., Cheng, L., Shapiro, D.N. & Winoto, A. Identification of human and mouse p19, a novel CDK4 and CDK6 inhibitor with homology to p16^{Ink4}. *Mol. Cell. Biol.* **15**, 2682–2688 (1995).
12. Ragione, F.D., Borriello, A., Giordani, L. & Iolascon, A. Cell division cycle alterations in human malignancies *Cancer J.* **10**, 151–156 (1997).
13. Pollock, P.M., Pearson, J.V. & Hayward, N.K. Compilation of somatic mutations of the CDKN2 gene in human cancers: non-random distribution of base substitutions. *Genes Chromosom. Cancer* **15**, 77–88 (1996).
14. Smith-Sorensen, B. & Hovig, E. CDK2A (p16^{INK4A}) somatic and germline mutations. *Human Mutation* **7**, 294–303 (1996).
15. Miller, C.W. *et al.* The p19^{INK4D} cyclin dependent kinase inhibitor gene is altered in osteosarcoma. *Oncogene* **15**, 231–235 (1997).
16. Gemma, A., *et al.* Molecular analysis of the cyclin-dependent kinase inhibitor genes p15$^{INK4b/MTS2}$, p16$^{INK4/MTS1}$, p18 and p19 in human cancer cell lines. *Int. J. Cancer* **68**, 605–611 (1996).
17. Lux, S., John, K. & Bennett, V. Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure with homology to tissue-differentiation and cell-cycle control proteins. *Nature* **344**, 36–42 (1990).
18. Michaely, P. & Bennett, V. The ANK repeat: a ubiquitous motif involved in macromolecular recognition. *Trends Cell. Biol.* **2**, 127–130 (1992).
19. Bork, P. Hundreds of ankyrin-like repeats in functionally diverse proteins: mobile modules that cross phyla horizontally? *Proteins* **17**, 363–374 (1993).
20. Pabo, C.O. & Sauer, R.T. Transcription factors: structural families and principles of DNA recognition. *Annu. Rev. Biochem.* **61**, 1053–1095 (1992).
21. Pines, J. Cyclins and cyclin-dependent kinases: theme and variations. *Adv. Cancer Res.* **66**, 181–212 (1995).
22. Luh, F.Y. *et al.* Structure of the cyclin-dependent kinase inhibitor p19^{Ink4d}. *Nature* **389**, 999–1003 (1997).
23. Svetlana, G. & Pavletich, N.P. Structure of the p53 tumor suppressor bound to the ankyrin and SH3 domains of p53BP2. *Science* **274**, 1001–1005 (1996).
24. Tevelev, A., *et al.* Tumor suppressor p16^{INK4}: structural charcterization of wild-type and mutant proteins by NMR and circular dichroism. *Biochemistry* **35**, 9475–9487 (1996).
25. Kalus, W., *et al.* NMR structural characterization of the CDK inhibitor p19^{INK4d}. *FEBS Lett.* **401**, 127–132 (1997).
26. Yang, Y., Rao, S., Walker, E., Sen, S. & Qin, J. Nuclear magnetic resonance assignment and secondary structure of an ankyrin-like repeat-bearing protein: Myotrophin *Protein Sci.* **6**, 1347–1351 (1997).
27. Fahraeus, R., Paramio, J.M., Ball, K.L., Lain, S. & Lane, D.P. Inhibition of pRb phosphorylation and cell-cycle progression by a 20-residue peptide derived from p16$^{CDKN2/INK4}$. *Curr. Biol.* **6**, 84–91 (1996).
28. Coleman, K.G., *et al.* Identification of CDK4 sequences involved in cyclin D1 and p16 binding. *J. Biol. Chem.* **272**, 18869–18874 (1997).
29. Wolfel, T. *et al.* A p16^{INK4a}-insensitive CDK4 mutant targeted by Cytolytic T lymphocytes in a human melanoma. *Science* **269**, 1281–1284 (1995).
30. De Bondt, H.L. *et al.* Crystal structure of cyclin-dependent kinase 2. *Nature* **363**, 595–602 (1993).
31. Russo, A.A., Jeffrey, P.D., Patten, A.K., Massague, J. & Pavletich, N.P. Crystal structure of the p27Kip1 cyclin-dependent-kinase inhibitor bound to the cyclin A-CDK2 complex. *Nature* **382**, 325–331 (1996).
32. Neidhardt, F.C., Bloch, P.L. & Smith, D.F. Culture medium for enterobacteria. *J. Bacteriol.* **119**, 736–747 (1974).
33. Doublie, S. Preparation of selenomethionyl proteins for phase determination. *Meth. Enz.* 523–530 (1997).
34. Kabsch, W. Automatic indexing of rotation diffraction patterns. *J. Appl. Crystallogr.* **21**, 67–71 (1988).
35. Kabsch, W. Evaluation of single-crystal X-ray diffraction data from a position-sensitive detector. *J. Appl. Crystallogr.* **21**, 916–924 (1988).
36. Gewirth, D., Otwinowski, Z. & Minor, W. *The HKL Version 1.0 manual* (Yale University Press, New Haven, Connecticut; 1993).
37. Furey, W. & Swaminathan, S. PHASES: A program package for the processing and analysis of diffraction data from macromolecules. *Meth. Enz.* **277**, 590–620 (1995).
38. Jones, T.A. A graphics model building and refinement system for macromolecules. *J. Appl. Crystallogr.* **11**, 268–272 (1978).
39. Jones, T.A. & Kjeldguard, M. *Manual for O, version 6.1* (Aarhus University Press, Uppsula, Sweden; 1996).
40. Brunger, A.T. The free R value: a novel statistical quantity assessing the accuracy of crystal structures. *Nature* **335**, 472–474 (1992).
41. Brunger, A.T. *X-PLOR Manual, Version 3.8.* (Yale University Press, New Haven, Connecticut; 1996).
42. Brunger, A.T., Kuriyan, J. & Karplus, M. Crystallographic R factor refinement by molecular dynamics. *Science* **235**, 458–460 (1987).
43. Jiang, J.S. & Brunger, A.T. Protein hydration observed by X-ray diffraction: solvation properties of penicillopepsin and neuraminidase crystal structures. *J. Mol. Biol.* **243**, 100–115 (1994).
44. Hodel, A., Kim, S.-H. & Brunger, A.T. Model bias in macromolecular crystal structures. *Acta Crystallogr.* **A48**, 851–858 (1992).
45. Foulkes, W.D., Flanders, T.Y., Pollock, P.M. & Hayward, N.K. The CDKN2A (p16) gene and human cancer. *Mol. Medicine* **3**, 5–20 (1997).
46. Kraulis, P.J. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
47. Merritt, E.A. & Murphy, M.E.P. RASTER3D version 2.0: a program for photorealistic molecular graphics. *Acta Crystallogr.* **D50**, 869–873 (1994).
48. Nicholls, A., Sharp, K. & Honig, B. Protein folding ans association: insights from interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281–296 (1991).

REPRINTED FROM:

17th International Cancer Congress

Rio de Janeiro (Brazil), August 24-28, 1998

Ed
MA OS
RICA O LENTIANI
RUY BEVILACQUA

M ND ZZI EDITOR
INTERNATIONAL PROCEEDINGS DIVISION

17th International
CANCER
CONGRESS

Rio de Janeiro, Brazil
24-28 August 1998

Structure of the p18^{INK4c} protein: insights into ankyrin-like repeat structure/function and CDK4/6 inhibition by the p16^{INK4} tumor suppressor

R. MARMORSTEIN, R. VENKATARAMANI* and K. SWAMINATHAN**

The Wistar Institute, the Department of Chemistry and the Department of Biochemistry and Biophysics University of Pennsylvania, Philadelphia, PA (USA)
** The Wistar Institute and the Department of Biochemistry and Biophysics University of Pennsylvania, Philadelphia, PA (USA)*
*** The Wistar Institute, Philadelphia, PA (USA)*

SUMMARY

p18^{INK4c} is a member of a family of INK4 proteins that function to arrest the G1 to S cell cycle transition by inhibiting the activity of the cyclin-dependent kinases (CDK) 4 and 6. The X-ray crystal structure of the human p18^{INK4c} protein reveals an elongated molecule comprised of 5 contiguous 32-33 residue ankyrin-like repeat units containing a mixed α/β fold. Comparison with the functionally unrelated ankyrin-repeat bearing regions of the 53BP2 and GABPβ proteins shows a highly homologous scaffold while subtle differences between the three proteins appear to be correlated with their respective protein-specific functions. A mapping of INK4 conserved residues and tumor-derived p16^{INK4} mutations onto the p18^{INK4c} structure suggests that conserved and mutationally sensitive residues play important roles in protein stability with a subset of these residues mapping to a putative CDK4/6 binding surface.

INTRODUCTION

Progression through the eukaryotic cell cycle correlates with the activity of three protein families cyclin-dependent kinases (CDKs), cyclins and CDK inhibitory proteins (CKIs). A major control point for the cell cycle occurs during the G1 to S transition during which time the CDK4/cyclinD1 and CDK6/cyclinD1 complexes phosphorylate the retinoblastoma gene product, pRb. Upon phosphorylation, pRb stimulates progression into the S phase of the cell cycle by allowing E2F transcription factors to activate genes required for S phase progression. The INK4 family of proteins, that includes $p16^{INK4}$, $p15^{INK4b}$, $p18^{INK4c}$ and $p19^{INK4d}$,play a key role in mediating S phase progression by directly inhibiting the kinase activity of specific CDK/cyclin complexes. The MTS1 gene encoding $p16^{INK4}$ has received considerable attention since it is deleted, silenced or mutated in a variety of tumor types, including pancreatic adenocarcinoma, esophageal carcinomas, biliary tract cancers, familial melanoma and pancreatic cancer (Ragione, et al., 1997). The homology among the INK4 proteins is restricted to a region containing four 32- 33 amino acid ankyrin-like repeats shared in more than 150 proteins, found from yeast to man, and occurring in functionally diverse proteins such as enzymes, toxins, and transcription factors (Bork, 1993).

MATERIALS AND METHODS

The bacterial overexpression, purification, crystallization and 1.95 Å structure determination of the full-length human $p18^{INK4c}$ protein is discussed elsewhere (Venkataramani, et al., 1998). The structure was determined by multiple isomorphous replacement and refined to an working R factor of 21.1% (R_{free} = 29.1%) with excellent geometry ($RMS_{bond\ length}$ = 0.009 Å, $RMS_{bond\ angle}$ = 1.5°).

RESULTS AND CONCLUSIONS

Overall structure of the $p18^{INK4c}$ protein. The $p18^{INK4c}$ structure adopts an elongated shape and is comprised of 5 contiguous 32 or 33-amino acid ankyrin-like segments with similar sequence and structure (fig. 1). Each repeat unit is comprised of a β-sheet, helix-turn-helix, extended strand, β-sheet element. The helices and β-sheets are arranged in anti-parallel fashion, with the plane of the β-sheet regions aligned roughly orthogonal to the axis of the helical segments. Except for the N- and C-terminal repeat elements, each ankyrin-like repeat unit interacts with its N- and C-neighboring repeat units via helix-helix and β-sheet - β-sheet interactions, where the helices form helical bundles and the sheet regions form 6-residue antiparallel β-hairpins. Several conserved features characterize and facilitate the fold of the ankyrin-like repeat segments of the 18^{INK4c} structure (fig. 2A). Most importantly, Gly residues at positions 2 and 13 of each repeat facilitate tight turns between secondary structural elements (fig 1A). There are also extensive van der Waals interactions that mediate the association between neighboring ankyrin repeats (fig. 1A). Specifically, positions 5, 10, 17, 20, and 21 of each repeat have conserved hydrophobic residues that carry out conserved inter-repeat stabilizing functions (fig. 1A). The β-hairpins at the junctions of each of the repeat units interact with neighboring β-hairpins predominantly through salt bridges between residues that are generally not conserved. Taken together, the overall fold and elongated shape of the $p18^{INK4c}$ protein appears to depend on the interactions between neighboring ankyrin-repeat units.

Implications for ankyrin-like repeat structure in homologous proteins. Correlation of the ankyrin-like repeat consensus sequence with the $p18^{INK4c}$ structure reveals that residues facilitating the fold of the ankyrin-like repeat segments and mediating

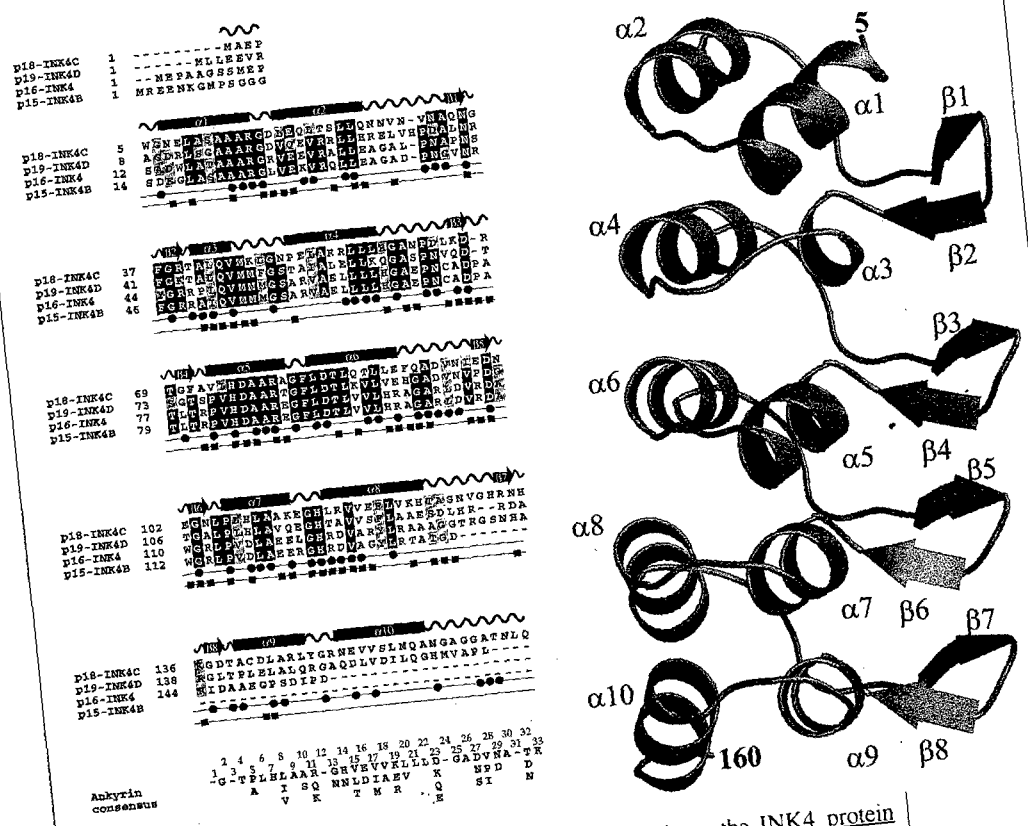


Fig. 1. Structure of the $p18^{INK4c}$ protein and its relationship to the INK4 protein family. (A) Sequence alignment of the INK4 protein family. Residues that are conserved in ankyrin repeat proteins in two or more homologs are indicated by a black filled circle, and $p16^{INK4}$ residues that are altered in primary tumors is indicated by a black square below the corresponding sequence position. The ankyrin consensus sequence is indicated below with a numbering scheme shown above the corresponding amino acid. (B) Overall structure of the $p18^{INK4c}$ protein.

interactions between neighboring repeat units are highly conserved in other ankyrin repeat domains (fig. 2A). This suggests that the ankyrin repeat unit maintains a conserved three-dimensional fold and that multiple units will align in a very homologous way in otherwise functionally unrelated ankyrin repeat-bearing proteins. A comparison of the $p18^{INK4c}$ structure with the X-ray crystal structures of the functionally unrelated ankyrin repeat regions from the 53BP2 (Svetlana and Pavletich, 1996) and GABPβ (Batchelor, et al., 1998) proteins confirms that functionally unrelated ankyrin-repeat domains have high structural homology (figs. 2B and 2C). Specifically, the RMS deviation between α-carbon atoms between these three proteins is less than 1.6 Å. This superposition excludes repeat unit 2 of $p18^{INK4c}$ and repeat unit 4 of 53BP2 which show interesting deviations that will be discussed below.

A detailed analysis of protein complexes with 53BP2 and GABPβ show that regions with the greatest deviations from the conserved ankyrin-repeat super-structure are important for protein-specific function. Specifically, 53BP2 shows a large structural deviation in the fourth repeat unit which makes directly contacts to

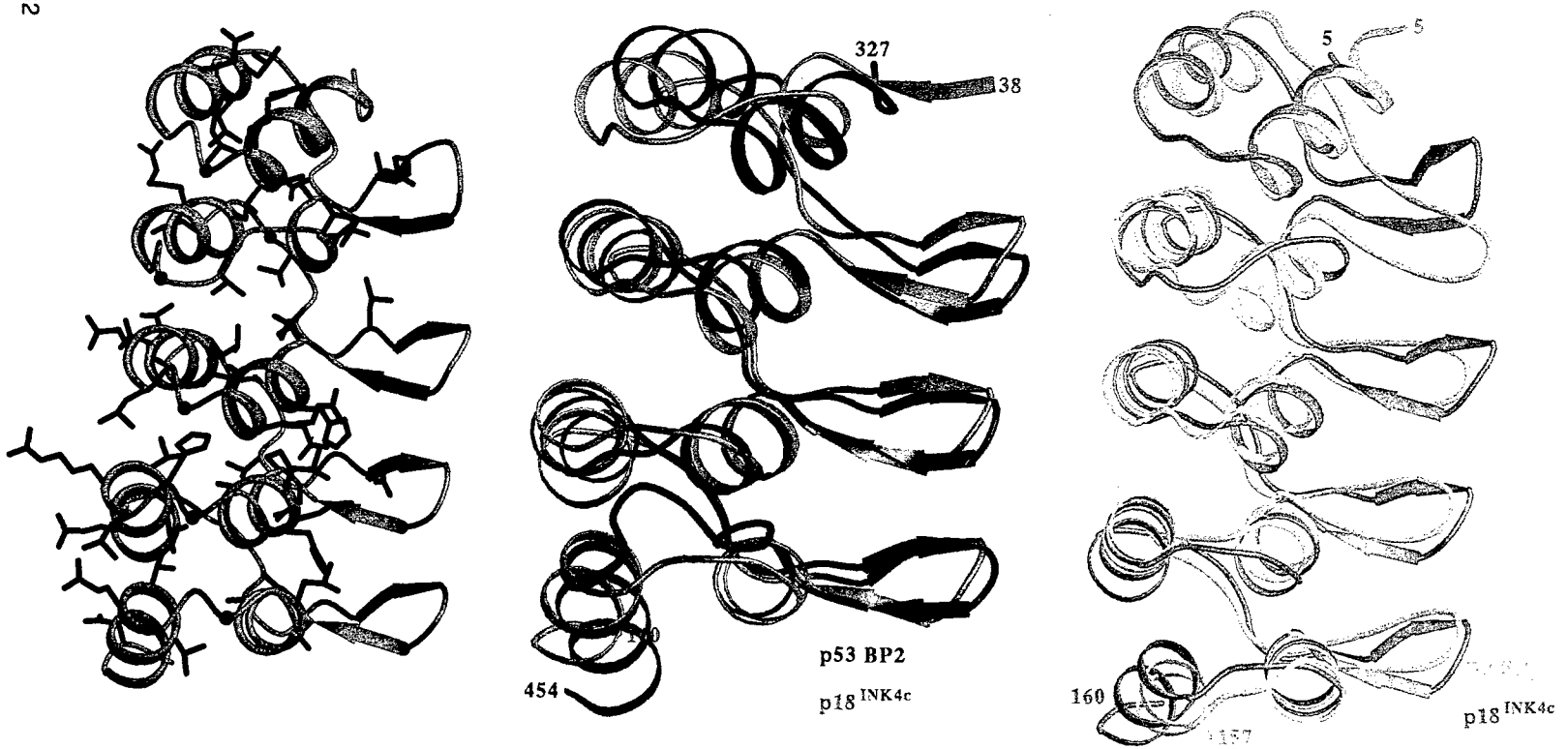


Fig. 2. Relationship of the $p18^{INK4c}$ structure to the structure of functionally unrelated ankyrin-like repeat regions. (A) Residues that are conserved among ankyrin-like domains (red side-chains) are mapped onto the $p18^{INK4c}$ structure, and glycines are shown as red balls. (B) Superposition between the $p18^{INK4c}$ (repeats 1-4) and 53BP2 proteins. (C) Superposition between the $p18^{INK4c}$ and GABPβ proteins.

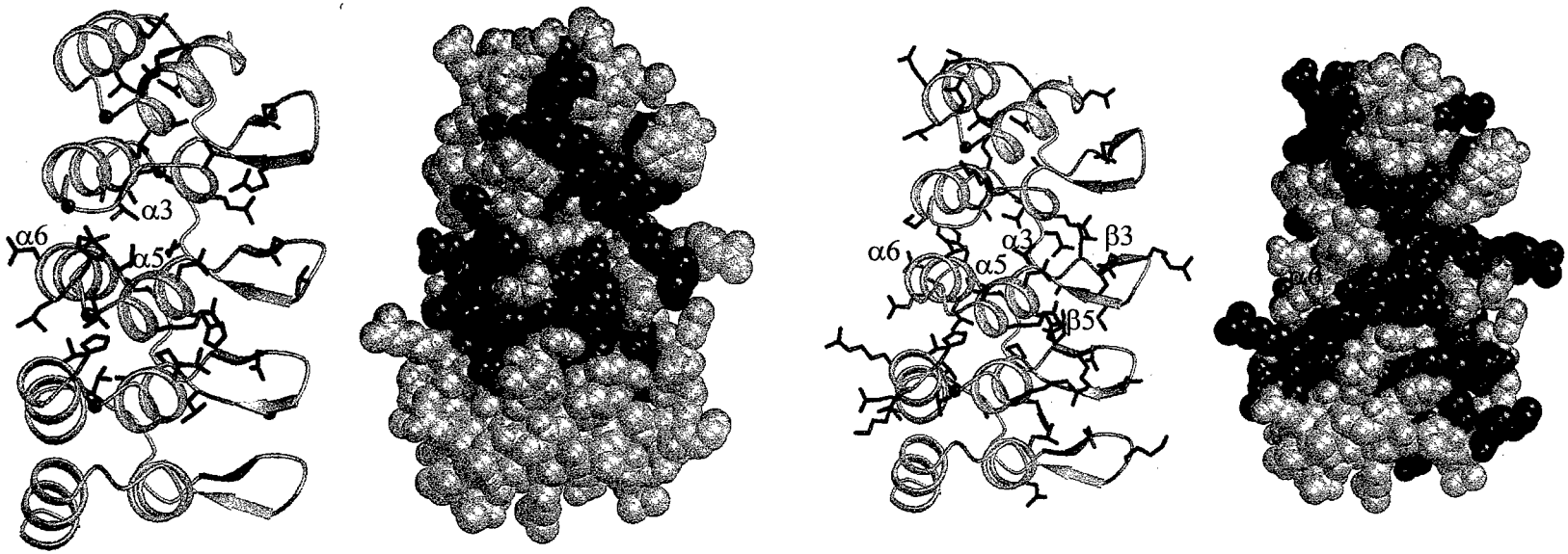


Fig. 3. Implications of the $p18^{INK4c}$ structure for INK4 proteins and tumor-derived mutations in the $p16^{INK4}$ protein. (A) Residues that are strictly conserved among INK4 proteins are highlighted in red; left, conserved residues are superimposed onto a schematic representation of the $p18^{INK4c}$ structure; right, a space filling model in the same orientation as the left view. (B) Residues of $p18^{INK4c}$ that correspond to positions of $p16^{INK4}$ tumor-derived mutations are highlighted in red. The views are as indicated in (A).

the p53 protein and is the site of linkage to an SH3 domain that also makes extensive contacts to the p53 protein (Svetlana and Pavletich, 1996) (fig. 2B). GABPβ shows the largest structural deviation in the β-hairpin region that mediates contact to the GABPα subunit of the heteromeric GABPα/β complex bound to DNA (Batchelor, et al., 1998) (fig. 2C). Taken together, regions of ankyrin-repeat proteins that deviated from the ankryrin-like repeat super-structure appear to play important roles for protein-specific function.

Implications for other INK4 proteins and for $p16^{INK4}$ tumor-derived mutations. Based on the assumption that INK4 proteins have very similar three-dimensional structures we mapped strictly conserved residues within the INK4 family and $p16^{INK4}$ tumor-derived mutations onto the $p18^{INK4c}$ structure (fig. 3). Most of these residues localize to positions involved in ankyrin repeat conformation or repeat-repeat interaction, suggesting that a large percentage of $p16^{INK4}$ missense mutations result in reduced $p16^{INK4}$ stability. Interestingly, strictly conserved and/or positions of tumor-derived $p16^{INK4}$ mutations that are solvent exposed are clustered around the α3 helix and the proceeding loop of ankyrin-like repeat unit 2, and the α5-turn-α6 region of ankyrin-like repeat unit 3 (fig. 3A) implicating these regions to play an important role in CDK4/6 inhibition. Moreover, the α3 helix of the second repeat is unusually short relative to the other helices of the structure (fig. 1A). This unusual feature of the α3 helix was also noted in the recently published three-dimensional solution structures of the $p19^{INK4d}$ (Luh, et al., 1997), and $p16^{INK4}$ (Byeon, et al., 1998) proteins, suggesting that the unusual structural feature of the α3 helix may have functional importance in INK4-mediated inhibition of CDK4/6.

The mutagenesis work of Coleman and coworkers (Coleman, et al., 1997) have suggested that INK4-mediated inhibition of CDK4 is mediated through the N-terminal lobe of the bi-lobed CDK4 structure (based on the homologous CDK2 structure). Taken together with our studies, this suggests that the α3-loop and α5-turn-α6 regions of INK4 proteins interact with the N-terminal lobe region of CDK4 and CDK6. Presumably, these interactions facilitate inhibition either through displacement of the cyclin D regulatory subunit, through direct inhibition of the CDK catalytic site located at the interface of the N- and C- terminal lobes, or through a combination of both mechanisms.

REFERENCES

RAGIONE, F. D., BORRIELLO, A., GIORDANI, L. et. al. Cell division cycle alterations in human malignancies, Cancer J. *10:* 151-156, 1997.

BORK, P. Hundreds of ankyrin-like repeats in functionally diverse proteins: mobile modules that cross phyla horizontally?, Proteins. *17:* 363-374, 1993.

VENKATARAMANI, R., SWAMINATHAN, K., AND MARMORSTEIN, R. Crystal structure of the CDK4/6 inhibitory protein $p18^{INK4c}$ provides insights into ankyrin-like repeat structure and function and tumor-derived $p16^{INK4}$ mutations., Nature Struc. Biol. *5:* 74-81, 1998.

SVETLANA, G. AND PAVLETICH, N. P. Structure of the p53 tumor suppressor bound to the ankyrin and SH3 domains of p53BP2., Science. *274:* 1001-1005, 1996.

BATCHELOR, A. H., PIPER, D. E., BROUSSE, F. C. D. L. et. al. The Structure of GABPα/β: An ETS domain-ankyrin repeat heterodimer bound to DNA., Science. *279:* 1037-1041, 1998.

LUH, F. Y., ARCHER, S. J., DOMAILLE, P. J. et. al. Structure of the cyclin-dependent kinase inhibitor $p19^{Ink4d}$., Nature. *389:* 999-1003, 1997.

BYEON, I. J. L., LI, J. N., ERICSON, K. et. al. Tumor suppressor p16(INK4A):

Determination of solution structure and analyses of its interaction with cyclin-dependent kinase 4, Molecular Cell. *1:* 421-431, 1998.
COLEMAN, K. G., WAUTLET, B. S., MORRISSEY, D. et. al. Identification of CDK4 sequences involved in cyclin D1 and p16 binding., J. Biol. Chem. *272:* 18869-18874, 1997.

MONDUZZI EDITORE

Via Ferrarese, 119/2 - 40128 Bologna, Italy

Tel ...39-51-370337 - Fax ...39-51-370529